

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 091980,862B  
Source: FWI  
Date Processed by STIC: 6/9/85

***ENTERED***



IFW16

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005  
TIME: 15:29:29

Input Set : A:\29029101.app  
Output Set: N:\CRF4\06092005\I980862B.raw

3 <110> APPLICANT: SCHLEHUBER, STEFFEN  
 5 <120> TITLE OF INVENTION: MUTEINS OF THE BILIN-BINDING PROTEIN  
 7 <130> FILE REFERENCE: 029029-0101  
 9 <140> CURRENT APPLICATION NUMBER: 09/980,862B  
 10 <141> CURRENT FILING DATE: 2002-02-27  
 12 <150> PRIOR APPLICATION NUMBER: PCT/DE00/01873  
 13 <151> PRIOR FILING DATE: 2000-06-08  
 15 <150> PRIOR APPLICATION NUMBER: DE 199 26 068.0  
 16 <151> PRIOR FILING DATE: 1999-06-08  
 18 <160> NUMBER OF SEQ ID NOS: 29  
 20 <170> SOFTWARE: PatentIn Ver. 2.1  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 1219  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Artificial Sequence  
 27 <220> FEATURE:  
 28 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 29 pBBP20 nucleic acid sequence  
 31 <220> FEATURE:  
 32 <221> NAME/KEY: CDS  
 33 <222> LOCATION: (22)..(1209)  
 35 <220> FEATURE:  
 36 <221> NAME/KEY: sig\_peptide  
 37 <222> LOCATION: (22)..(84)  
 39 <220> FEATURE:  
 40 <221> NAME/KEY: mat\_peptide  
 41 <222> LOCATION: (85)..(1209)  
 42 <223> OTHER INFORMATION: fusion protein of bilin-binding protein, Strep-tag II  
 43 and fragment of phage coat protein pIII  
 45 <220> FEATURE:  
 46 <221> NAME/KEY: misc\_feature  
 47 <222> LOCATION: (85)..(606)  
 48 <223> OTHER INFORMATION: mature bilin-binding protein  
 50 <220> FEATURE:  
 51 <221> NAME/KEY: misc\_feature  
 52 <222> LOCATION: (607)..(636)  
 53 <223> OTHER INFORMATION: Strep-tag II-affinity tag  
 55 <220> FEATURE:  
 56 <221> NAME/KEY: misc\_feature  
 57 <222> LOCATION: (637)..(639)  
 58 <223> OTHER INFORMATION: amber stop codon  
 60 <220> FEATURE:  
 61 <221> NAME/KEY: misc\_feature

(Pg.6)

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005

TIME: 15:29:29

Input Set : A:\29029101.app

Output Set: N:\CRF4\06092005\I980862B.raw

62 <222> LOCATION: (640)..(1209)  
 63 <223> OTHER INFORMATION: amino acids 217-406 of coat protein pIII  
 65 <400> SEQUENCE: 1

66	tctagttaac	gagggcaaaa	a	atg	aaa	aag	aca	gct	atc	gcg	att	gca	gtg	51			
67				Met	Lys	Lys	Thr	Ala	Ile	Ala	Ile	Ala	Val				
68				-20									-15				
70	gca	ctg	gct	ggt	ttc	gct	acc	gta	gcg	cag	gcc	gac	gtg	tac	cac	gac	99
71	Ala	Leu	Ala	Gly	Phe	Ala	Thr	Val	Ala	Gln	Ala	Asp	Val	Tyr	His	Asp	
72	-10			-5			-1	1						5			
74	ggt	gcc	tgt	ccc	gaa	gtc	aag	cca	gtc	gac	aac	ttc	gac	tgg	tcc	cag	147
75	Gly	Ala	Cys	Pro	Glu	Val	Lys	Pro	Val	Asp	Asn	Phe	Asp	Trp	Ser	Gln	
76				10				15						20			
78	tac	cat	gtt	aaa	tgg	tgg	gaa	gtc	gcc	aaa	tac	ccc	aac	tca	gtt	gag	195
79	Tyr	His	Gly	Lys	Trp	Trp	Glu	Val	Ala	Lys	Tyr	Pro	Asn	Ser	Val	Glu	
80				25			30			35							
82	aag	tac	gga	aag	tgc	gga	tgg	gct	gag	tac	act	cct	gaa	ggc	aag	agt	243
83	Lys	Tyr	Gly	Lys	Cys	Gly	Trp	Ala	Glu	Tyr	Thr	Pro	Glu	Gly	Lys	Ser	
84				40			45			50							
86	gtc	aaa	gtt	tcg	aac	tac	cac	gta	atc	cac	ggc	aag	gaa	tac	ttt	att	291
87	Val	Lys	Val	Ser	Asn	Tyr	His	Val	Ile	His	Gly	Lys	Glu	Tyr	Phe	Ile	
88				55			60			65							
90	gaa	gga	act	gcc	tac	cca	gtt	ggt	gac	tcc	aag	att	gga	aag	atc	tac	339
91	Glu	Gly	Thr	Ala	Tyr	Pro	Val	Gly	Asp	Ser	Lys	Ile	Gly	Lys	Ile	Tyr	
92	70			75				80						85			
94	cac	agc	ctg	act	tac	gga	ggt	gtc	acc	aag	gag	aac	gta	ttc	aac	gta	387
95	His	Ser	Leu	Thr	Tyr	Gly	Gly	Val	Thr	Lys	Glu	Asn	Val	Phe	Asn	Val	
96				90			95			100							
98	ctc	tcc	act	gac	aac	aag	aac	tac	atc	atc	gga	tac	tac	tgc	aaa	tac	435
99	Leu	Ser	Thr	Asp	Asn	Lys	Asn	Tyr	Ile	Ile	Gly	Tyr	Tyr	Cys	Lys	Tyr	
100				105			110			115							
102	gac	gag	gac	aag	aag	gga	cac	caa	gac	ttc	gtc	tgg	gtg	ctc	tcc	aga	483
103	Asp	Glu	Asp	Lys	Lys	Gly	His	Gln	Asp	Phe	Val	Trp	Val	Leu	Ser	Arg	
104				120			125			130							
106	agc	atg	gtc	ctt	act	ggt	gaa	gcc	aag	acc	gct	gtc	gag	aac	tac	ctt	531
107	Ser	Met	Val	Leu	Thr	Gly	Glu	Ala	Lys	Thr	Ala	Val	Glu	Asn	Tyr	Leu	
108				135			140			145							
110	atc	ggc	tcc	cca	gta	gtc	gac	tcc	cag	aaa	ctg	gta	tac	agt	gac	ttc	579
111	Ile	Gly	Ser	Pro	Val	Val	Asp	Ser	Gln	Lys	Leu	Val	Tyr	Ser	Asp	Phe	
112	150			155				160						165			
114	tct	gaa	gcc	gcc	tgc	aag	gtc	aac	aat	agc	aac	tgg	tct	cac	ccg	cag	627
115	Ser	Glu	Ala	Ala	Cys	Lys	Val	Asn	Asn	Ser	Asn	Trp	Ser	His	Pro	Gln	
116				170			175			180							
118	ttc	gaa	aaa	tag	gtc	ggc	ggc	ggc	tct	ggt	ggt	tct	ggc	ggc	ggc	ggc	675
119	Phe	Glu	Lys	Gln	Ala	Gly	Gly	Ser	Gly	Gly	Ser	Gly	Gly	Ser	Gly	Gly	
120				185			190			195							
122	tct	gag	ggt	ggt	ggc	tct	gag	ggt	ggc	ggt	tct	gag	ggt	ggc	ggc	tct	723
123	Ser	Glu	Gly	Gly	Ser	Glu	Gly	Gly	Ser	Glu	Gly	Gly	Ser	Gly	Gly	Ser	
124				200			205			210							
126	gag	gga	ggc	ggt	tcc	ggt	ggt	ggc	tct	ggt	tcc	ggt	gat	ttt	gat	tat	771

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005

TIME: 15:29:29

Input Set : A:\29029101.app  
 Output Set: N:\CRF4\06092005\I980862B.raw

127	Glu	Gly	Gly	Gly	Ser	Gly	Gly	Ser	Gly	Asp	Phe	Asp	Tyr					
128	215				220				225									
130	gaa	aag	atg	gca	aac	gct	aat	aag	ggg	gct	atg	acc	gaa	aat	gcc	gat	819	
131	Glu	Lys	Met	Ala	Asn	Ala	Asn	Lys	Gly	Ala	Met	Thr	Glu	Asn	Ala	Asp		
132	230				235				240				245					
134	gaa	aac	gcg	cta	cag	tct	gac	gct	aaa	ggc	aaa	ctt	gat	tct	gtc	gct	867	
135	Glu	Asn	Ala	Leu	Gln	Ser	Asp	Ala	Lys	Gly	Lys	Leu	Asp	Ser	Val	Ala		
136	250				255				255				260					
138	act	gat	tac	ggt	gct	gct	atc	gat	ggt	ttc	att	ggt	gac	gtt	tcc	ggc	915	
139	Thr	Asp	Tyr	Gly	Ala	Ala	Ile	Asp	Gly	Phe	Ile	Gly	Asp	Val	Ser	Gly		
140	265				270				275									
142	ctt	gct	aat	ggt	aat	ggt	gct	act	ggt	gat	ttt	gct	ggc	tct	aat	tcc	963	
143	Leu	Ala	Asn	Gly	Asn	Gly	Ala	Thr	Gly	Asp	Phe	Ala	Gly	Ser	Asn	Ser		
144	280				285				290									
146	caa	atg	gct	caa	gtc	ggt	gac	ggt	gat	aat	tca	cct	tta	atg	aat	aat	1011	
147	Gln	Met	Ala	Gln	Val	Gly	Asp	Gly	Asp	Asn	Ser	Pro	Leu	Met	Asn	Asn		
148	295				300				305									
150	ttc	cgt	caa	tat	tta	cct	tcc	ctc	cct	caa	tcg	gtt	gaa	tgt	cgc	cct	1059	
151	Phe	Arg	Gln	Tyr	Leu	Pro	Ser	Leu	Pro	Gln	Ser	Val	Glu	Cys	Arg	Pro		
152	310				315				320				325					
154	ttt	gtc	ttt	ggc	gct	ggt	aaa	cca	tat	gaa	ttt	tct	att	gat	tgt	gac	1107	
155	Phe	Val	Phe	Gly	Ala	Gly	Lys	Pro	Tyr	Glu	Phe	Ser	Ile	Asp	Cys	Asp		
156	330				335				335				340					
158	aaa	ata	aac	tta	ttc	cgt	ggt	gtc	ttt	gct	ttt	ctt	tta	tat	gtt	gcc	1155	
159	Lys	Ile	Asn	Leu	Phe	Arg	Gly	Val	Phe	Ala	Phe	Leu	Leu	Tyr	Val	Ala		
160	345				350				355									
162	acc	ttt	atg	tat	gta	ttt	tct	acg	ttt	gct	aac	ata	ctg	cgt	aat	aag	1203	
163	Thr	Phe	Met	Tyr	Val	Phe	Ser	Thr	Phe	Ala	Asn	Ile	Leu	Arg	Asn	Lys		
164	360				365				370									
166	gag	tct	taataagtt														1219	
167	Glu	Ser																
168	375																	
171	<210>	SEQ ID NO:	2															
172	<211>	LENGTH:	64															
173	<212>	TYPE:	DNA															
174	<213>	ORGANISM:	Artificial Sequence															
176	<220>	FEATURE:																
177	<223>	OTHER INFORMATION:	Description of Artificial Sequence:	Synthetic primer														
179	<220>	FEATURE:																
180	<221>	NAME/KEY:	modified_base															
181	<222>	LOCATION:	(35)..(36)															
182	<223>	OTHER INFORMATION:	a, t, c, g, other or unknown															
184	<220>	FEATURE:																
185	<221>	NAME/KEY:	modified_base															
186	<222>	LOCATION:	(38)															
187	<223>	OTHER INFORMATION:	a, t, c, g, other or unknown															
189	<220>	FEATURE:																
190	<221>	NAME/KEY:	modified_base															
191	<222>	LOCATION:	(41)..(42)															

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005

TIME: 15:29:29

Input Set : A:\29029101.app  
Output Set: N:\CRF4\06092005\I980862B.raw

192 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
194 <220> FEATURE:

195 <221> NAME/KEY: modified base  
196 <222> LOCATION: (44)..(45)

197 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
199 <400> SEQUENCE: 2

W--> 200 ccatggtaaa tggtgggaag tcgccaata ccccnknms nnsnnkaagt acggaaagtg 60  
201 cgga 64  
204 <210> SEQ ID NO: 3  
205 <211> LENGTH: 71  
206 <212> TYPE: DNA  
207 <213> ORGANISM: Artificial Sequence  
209 <220> FEATURE:  
210 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer  
212 <220> FEATURE:  
213 <221> NAME/KEY: modified base  
214 <222> LOCATION: (19)..(20)  
215 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
217 <220> FEATURE:  
218 <221> NAME/KEY: modified base  
219 <222> LOCATION: (46)..(47)  
220 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
222 <220> FEATURE:  
223 <221> NAME/KEY: modified base  
224 <222> LOCATION: (52)..(53)  
225 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
227 <400> SEQUENCE: 3

W--> 228 gggtaggcgg tacttcsnn aaagtattcc ttgccgtgga ttacmnnngta snncaaact 60  
229 ttgacactct t 71  
232 <210> SEQ ID NO: 4  
233 <211> LENGTH: 74  
234 <212> TYPE: DNA  
235 <213> ORGANISM: Artificial Sequence  
237 <220> FEATURE:  
238 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer  
240 <220> FEATURE:  
241 <221> NAME/KEY: modified base  
242 <222> LOCATION: (27)..(28)  
243 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
245 <220> FEATURE:  
246 <221> NAME/KEY: modified base  
247 <222> LOCATION: (33)..(34)  
248 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
250 <220> FEATURE:  
251 <221> NAME/KEY: modified base  
252 <222> LOCATION: (42)..(43)  
253 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
255 <220> FEATURE:  
256 <221> NAME/KEY: modified base

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005  
TIME: 15:29:29

Input Set : A:\29029101.app  
Output Set: N:\CRF4\06092005\I980862B.raw

257 <222> LOCATION: (54)..(55)  
 258 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 260 <400> SEQUENCE: 4  
**W--> 261 ccaagattgg aaagatctac cacagcnnsa ctnnkggagg tnnsaccvvs gagnnkgtat 60**  
 262 tcaacgtact ctcc 74  
 265 <210> SEQ ID NO: 5  
 266 <211> LENGTH: 78  
 267 <212> TYPE: DNA  
 268 <213> ORGANISM: Artificial Sequence  
 270 <220> FEATURE:  
 271 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer  
 273 <220> FEATURE:  
 274 <221> NAME/KEY: modified\_base  
 275 <222> LOCATION: (20)..(21)  
 276 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 278 <220> FEATURE:  
 279 <221> NAME/KEY: modified\_base  
 280 <222> LOCATION: (26)..(27)  
 281 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 283 <220> FEATURE:  
 284 <221> NAME/KEY: modified\_base  
 285 <222> LOCATION: (53)..(54)  
 286 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 288 <220> FEATURE:  
 289 <221> NAME/KEY: modified\_base  
 290 <222> LOCATION: (59)..(60)  
 291 <223> OTHER INFORMATION: a, t, c, g, other or unknown  
 293 <400> SEQUENCE: 5  
**W--> 294 tctggagaga acccagacmn ngtcsnnngtg tcccttcttg tcctcgctgt asnngcamnn 60**  
 295 gtatccgatg atgttagtt 78  
 298 <210> SEQ ID NO: 6  
 299 <211> LENGTH: 36  
 300 <212> TYPE: DNA  
 301 <213> ORGANISM: Artificial Sequence  
 303 <220> FEATURE:  
 304 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer  
 306 <400> SEQUENCE: 6  
 307 cttegactgg tcccagtacc atggtaaatg gtggga 36  
 310 <210> SEQ ID NO: 7  
 311 <211> LENGTH: 37  
 312 <212> TYPE: DNA  
 313 <213> ORGANISM: Artificial Sequence  
 315 <220> FEATURE:  
 316 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer  
 318 <400> SEQUENCE: 7  
 319 caccagtaag gaccatgctt ctggagagca cccagac 37  
 322 <210> SEQ ID NO: 8  
 323 <211> LENGTH: 46  
 324 <212> TYPE: DNA

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005  
TIME: 15:29:30

Input Set : A:\29029101.app  
Output Set: N:\CRF4\06092005\I980862B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 36,39,38,41,42,44,45  
Seq#:3; N Pos. 19,20,46,47,52,53  
Seq#:4; N Pos. 27,28,39,34,47,48,54,55  
Seq#:5; N Pos. 20,21,26,27,53,54,55,60  
Seq#:13; N Pos. 29,30,38,39,47,48,50,51,53,54,56,57

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:16; Line(s) 906

**VERIFICATION SUMMARY**  
PATENT APPLICATION: US/09/980,862B

DATE: 06/09/2005  
TIME: 15:29:30

Input Set : A:\29029101.app  
Output Set: N:\CRF4\06092005\I980862B.raw

L:200 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0  
L:228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0  
L:261 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0  
L:294 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0  
L:569 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0  
L:636 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 14, CDS LOCATION:22..1209  
L:874 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:16  
L:878 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:16